

Sintetizador Paramétrico de Lengua de Signos Española*

Spanish Sign Language Parametric Synthesizer

Fernando López-Colino

Departamento de Ingeniería Informática
Universidad Autónoma de Madrid. Fco. Tomás y Valiente 11
Madrid, 28049, Spain
fj.lopez@uam.es

Resumen: Este trabajo presenta los resultados de la tesis doctoral realizada por Fernando Jesús López Colino y dirigida por el Dr. José Colás Pasamontes. Esta tesis fue defendida el 15 de octubre de 2009 y calificada con Sobresaliente Cum Laude por unanimidad. La principal novedad lingüística aportada por la tesis es el uso del modelo fonológico propuesto por Liddell y Johnson en la síntesis de lengua de signos española (LSE). El modelo requiere la descripción del aspecto temporal de la LS, que no es contemplado por las notaciones existentes. Por ese motivo, la descripción fonémica de los signos y la evolución temporal se recogen en una base de datos relacional. En las evaluaciones de inteligibilidad con usuarios nativos de LSE se ha obtenido una tasa de reconocimiento media del 97 % para signos aislados, del 95 % para frases completas y del 85 % para construcciones clasificatorias.

Palabras clave: Lengua de signos española, Síntesis automática, Discapacitados auditivos, Construcciones clasificatorias

Abstract: PhD Thesis written by Fernando Jesús López Colino under the supervision of Dr. José Colás Pasamontes. The author was examined in October 15th, 2009 and the grade obtained was *Sobresaliente Cum Laude* by unanimous decision. The Thesis' main linguistic contribution is to use the phonological model proposed by Liddell and Johnson applied to the Spanish Sign Language (LSE) synthesis. This model requires defining the temporal aspect of the Sign Language (LS), which is not described by existing notations. Therefore, the phonologic description and the temporal aspects are stored in a relational database. The intelligibility evaluations, using LSE natives, showed the following recognition rates: 97 % for isolated signs, 95 % for complete sentences and 85 % for classifier constructions.

Keywords: Spanish Sign Language, Automatic Synthesis, Deaf People, Classifier Constructions

1. Introducción

La síntesis de lengua de signos (LS) es un área de investigación con una antigüedad de apenas una década, periodo en el cual han surgido numerosas soluciones. La técnica más utilizada es la síntesis paramétrica, donde se parte de una descripción de los signos en función de los siete parámetros formativos quinésicos (PFQ) que los conforman (también conocidos como fonemas). Aunque esta técnica no produce los resultados más naturales (como los obtenidos por sistemas de animación manual (Huenerfauth, 2009)), sí es

la única que ofrece la flexibilidad y grado de control necesarios para gestionar la complejidad de la LS, razón por la que ha sido elegida para este trabajo. Este trabajo realiza la primera aproximación genérica a la síntesis de lengua de signos española (LSE).

2. Estado del Arte

Los sintetizadores paramétricos actuales (Zwitterslood et al., 2004; Elliott et al., 2008; Kennaway, Glauert, y Zwitterlood, 2007) utilizan una adaptación a XML de las notaciones iconográficas que describen los PFQ de los signos a sintetizar (Prillwitz et al., 1989). Esta aproximación es equivalente a la síntesis de voz, en la que se utiliza una notación gráfica para representar los fonemas. El uso de esta aproximación presenta ciertos incon-

* Agradecimientos a los colaboradores de la Fundación Confederación Nacional de Sordos de España (FCNSE) y a Almudena Sánchez por los conocimientos en LSE aportados; al programa FPU-UAM por su financiación económica.

venientes, tales como que la descripción de un mensaje a sintetizar es una tarea compleja o la falta de definición en la dimensión temporal de los signos: cada uno de los PFQ se definen como una secuencia de unidades en las que es necesario definir el instante en que se representan cada una de ellas; de igual modo, la síntesis de la LS requiere de la definición de sincronización entre dichas secuencias.

3. Desarrollo Realizado

En este trabajo afrontamos la síntesis paramétrica de LS desde una nueva perspectiva, siguiendo el modelo lingüístico propuesto en (Liddell y Johnson, 1989). Dicho modelo requiere mayor precisión en la definición fonémica de los signos, junto con la descripción del aspecto temporal de los mismos. Por este motivo, hemos optado por el uso de una base de datos relacional para el almacenamiento de dichas definiciones. Esta base de datos no sólo permite disponer de descripciones más precisas, detalladas y flexibles, sino que además permite definir representaciones paralelas, tanto de los propios signos como de los PFQ. Gracias a estas definiciones paralelas se puede contemplar en la misma base de datos la representación de los signos en distintos idiomas o dialectos y afrontar la síntesis de variantes emocionales para los mismos.

Una mayor flexibilidad de las descripciones de los signos, permite que los expertos en este modo de comunicación, especialmente signantes nativos, participen tanto en la evaluación como en la propia definición de los signos. Para ello se ha diseñado un conjunto de aplicaciones específicas para esta tarea, basadas en una interfaz visual. Dado que la notación de entrada ya no debe describir los signos mediante los PFQ, podemos definir una notación de entrada en un nivel superior de abstracción, lo que facilita la descripción de mensajes en LS. Esta nueva notación, que hemos denominado HLSML, contempla las tres categorías semánticas de la LS: el diccionario dactilológico, los signos establecidos y las construcciones clasificatorias. Aunque la notación HLSML se centra en la descripción sintáctica del mensaje signado, permite la descripción fonémica de los signos del mensaje, cuando éstos no están contemplados en la base de datos.

Otro aspecto novedoso a destacar del sintetizador desarrollado es el diseño del esqueleto del avatar. El avatar utilizado inclu-

ye características especiales para la gestión fonémica independiente de la LS.

4. Resultados Obtenidos

Los contenidos generados por el sintetizador han sido evaluados por 11 personas sordas, entre ellas expertos de la FCNSE. La evaluación se compuso de distintos test centrados en medir la inteligibilidad de signos aislados, frases completas y construcciones clasificatorias en LSE. El porcentaje de aciertos para los signos aislados fue de un 96 %, permitiendo respuesta libre, y de un 98 %, cuando se proponían cinco opciones posibles. En el caso de frases compuestas por secuencias de signos, se obtuvo una tasa de aciertos del 95 %. Por último, la evaluación de las construcciones clasificatorias obtuvo una tasa de aciertos del 85 %.

Bibliografía

- Elliott, Ralph, John Glauert, Richard Kennaway, Ian Marshal, y Éva Sáfár. 2008. Linguistic modelling and language-processing technologies for avatar-based sign language presentation. *Universal Access in the Information Society*, 6(4):375–391.
- Huenerfauth, Matt. 2009. A linguistically motivated model for speed and pausing in animations of american sign language. *ACM Transactions on Accessible Computing*, 2(2):1–31.
- Kennaway, Richard, John Glauert, y Inge Zwitterlood. 2007. Providing signed content on the internet by synthesized animation. *ACM Transactions on Computer-Human Interaction*, 14(15):1–29.
- Liddell, Scott K. y Robert E. Johnson. 1989. American sign language: The phonological base. *Sign Language Studies*, 64:195–278, fall.
- Prillwitz, Siegmund, Regina Leven, Heiko Zienert, Thomas Hanke, y Jan Herming. 1989. *HamNoSys. Version 2.0; Hamburg Notation System for Sign Languages. An introductory guide*. Signum-Verlag.
- Zwitterslood, Inge, Margriet Verlinden, Johan Ros, y Sanny van der Schoot. 2004. Synthetic signing for the deaf: esign. En *Proceedings of the Conference and Workshop on Assistive Technologies for Vision and Hearing Impairment*, Granada, Spain, Junio. CVHI.